

DOI:10.13718/j.cnki.xsxb.2019.01.016

基于增强多流形学习的监控视频追踪算法^①

李建新

东莞职业技术学院 计算机工程系, 广东 东莞 523808

摘要: 提出了一种多流形局部线性嵌入的流形学习算法, 为每个类的流形学习过程设计了一种监督的近邻点选择方法, 将流形-流形距离作为度量指标, 搜索最优的低维空间. 在视频追踪算法中对外部数据库进行图像训练预处理, 为 лица 检测建立级联分类器, 利用均值粒子滤波器结合跟踪校正策略对人脸图像实时跟踪, 采用多流形训练的结果从视频流的人脸集中检测出追踪的目标人脸. 仿真实验结果表明本算法对不同的数据集均获得了较高的检测率与较高的计算效率.

关键词: 流形学习; 局部线性嵌入; 视频监控; 目标检测; 目标追踪; 降低维度

中图分类号: TP391

文献标志码: A

文章编号: 1000-5471(2019)01-0095-07

目前 LLE(局部线性嵌入)算法^[1-2] 广泛地应用于图像分类与目标识别问题中^[3], 但经典的 LLE 算法难以学习多个流形^[4-8]. 本文在文献^[9]的 LLE 算法的基础上, 设计了基于 LLE 的多流形学习方法, 通过局部非线性、多流形的方法学习并保留数据集的类结构. 假设一个向量数据集为 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$, ϵ 是一个正实数, k 是一个正整数, $N(\mathbf{x}_i)$ 表示 \mathbf{x}_i 的近邻点, 包含了与 \mathbf{x}_i 距离最近的 k 个向量, 即: 对于每个 $\mathbf{x}_j \in N(\mathbf{x}_i)$, 其欧式距离 $\|\mathbf{x}_i - \mathbf{x}_j\|$ 小于 ϵ , 将数据样本 \mathbf{x}_i 线性地表示为 $N(\mathbf{x}_i)$ 中的近邻点. 通过最小化(1)式计算权重值 $\{\omega_{ij}\}_{\mathbf{x}_j \in N(\mathbf{x}_i)}$:

$$\epsilon = \mathbf{E}_1^{(i)}(\{\omega_{ij}\}_{\mathbf{x}_j \in N(\mathbf{x}_i)}) = \left[\mathbf{x}_i - \sum_{\mathbf{x}_j \in N(\mathbf{x}_i)} \omega_{ij} \mathbf{x}_j \right]^2 \quad (1)$$

约束条件为 $\sum_{\mathbf{x}_j \in N(\mathbf{x}_i)} \omega_{ij} = 1$, $\omega_{ij} > 0$; 如果 $\mathbf{x}_j \notin N(\mathbf{x}_i)$, 则 $\omega_{ij} = 0$. 可将(1)式转换为: $\epsilon^2 = \left[\mathbf{x}_i - \sum_{j=1}^k \omega_{ij} \mathbf{x}_j \right]^2 = \left[\sum_{j=1}^k \omega_{ij} (\mathbf{x}_i - \mathbf{x}_j) \right]^2 = \sum_{jk} \omega_{ij} \omega_{ik} \mathbf{G}_{jk}$, 式中 $\mathbf{G}_{jk} = (\mathbf{x}_i - \mathbf{x}_j)^\top (\mathbf{x}_i - \mathbf{x}_k)$ 是局部 Gram 矩阵,

其中 $\mathbf{x}_j, \mathbf{x}_k \in N(\mathbf{x}_i)$. 使用拉格朗日乘法计算重构权重 $\omega_{ij} = \frac{\sum_k \mathbf{G}_{jk}^{-1}}{\sum_{lm} \mathbf{G}_{lm}^{-1}}$, 使用重构权重搜索 d 维的向量 $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n \in \mathbb{R}^d$, 这些向量定义了 \mathbb{R}^d 中原数据集的 d 维嵌入, 其中 $d < D$, 通过最小化(2)式可获得 d 维的向量.

$$\mathbf{E}_2(\mathbf{Y}) = \mathbf{E}_2(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n) = \sum_{i=1}^n \left[\mathbf{y}_i - \sum_{\mathbf{y}_j \in N(\mathbf{y}_i)} \omega_{ij} \mathbf{y}_j \right]^2 \quad (2)$$

其中

① 收稿日期: 2017-01-18

基金项目: 广东省教育厅青年创新人才类项目(2017GkQNCX116, 2017GkQNCX119); 东莞市社会科技发展项目(2017507156388); 2018 东莞职业技术学院校企合作项目(政 201805).

作者简介: 李建新(1984-), 男, 讲师, 硕士, 主要从事模式识别、图形图像处理、智能软计算研究.

$$Y = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_n \end{bmatrix}_{n \times d} = \begin{bmatrix} \mathbf{y}_{11} & \mathbf{y}_{12} & \cdots & \mathbf{y}_{1d} \\ \mathbf{y}_{21} & \mathbf{y}_{22} & \cdots & \mathbf{y}_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{y}_{n1} & \mathbf{y}_{n2} & \cdots & \mathbf{y}_{nd} \end{bmatrix}_{n \times d} = [\mathbf{Y}_1 \mathbf{Y}_2 \cdots \mathbf{Y}_d]_{n \times d} \quad (3)$$

将(2)式改写为以下的矩阵形式:

$$E_2(\mathbf{Y}) = \|\mathbf{Y} - \mathbf{W}\mathbf{Y}\|^2 = \sum_{k=1}^d [\mathbf{Y}_k - \mathbf{W}\mathbf{Y}_k]^2 = \sum_{k=1}^d (\mathbf{Y}_k - \mathbf{W}\mathbf{Y}_k)^T (\mathbf{Y}_k - \mathbf{W}\mathbf{Y}_k) = \sum_{k=1}^d \mathbf{Y}_k^T (\mathbf{I} - \mathbf{W})^T (\mathbf{I} - \mathbf{W}) \mathbf{Y}_k = \sum_{k=1}^d \mathbf{Y}_k^T \mathbf{m} \mathbf{Y}_k$$

式中: \mathbf{W} 是 $n \times n$ 的权重矩阵, $\mathbf{m}_{N \times N} = (\mathbf{I} - \mathbf{W}_{n \times n})^T (\mathbf{I} - \mathbf{W}_{n \times n})$, \mathbf{Y}_i 均正交. 使用拉格朗日乘法计算最优解:

$$E_2(\mathbf{Y}) = \sum_{k=1}^d \mathbf{Y}_k^T \mathbf{m} \mathbf{Y}_k - \sum_{k=1}^d \lambda_k \left(\frac{1}{N} \mathbf{Y}_k^T \mathbf{Y}_k - 1 \right) \quad (4)$$

用(4)式对 $\mathbf{Y}_k (k=1, 2, \dots, d)$ 求导, $\mathbf{Y}_k^T \mathbf{m} - \lambda_k \mathbf{Y}_k^T = 0$ 则说明了 $\mathbf{m} \mathbf{Y}_k = \lambda_k \mathbf{Y}_k$, $\{\lambda_k\}$ 为 \mathbf{m} 的特征值, $\{\mathbf{Y}_k\}$ 则是对应的特征向量. 通过搜索 \mathbf{m} 底部的特征向量可将 $E_2(\mathbf{Y})$ 最小化. \mathbf{Y}_0^T 的一个特征值为 $\lambda_0 = 0$, 所以应当获得 \mathbf{m} 底部的 $d+1$ 个特征值及其特征向量, 然后忽略 \mathbf{Y}_0 .

1 系统架构

图 1 所示是本文系统的总体结构: 首先对外部数据库进行图像训练预处理, 为人脸检测建立级联分类器; 其次利用摄像头采集监控的视频信息, 通过级联分类器检测视频流的人脸图像; 然后利用均值粒子滤波器结合跟踪校正策略对人脸图像实时跟踪^[10]; 接着将追踪的目标人脸样本集作为训练库进行多流形局部嵌入的学习; 最后采用多流形训练的结果从视频流的人脸集中检测出追踪的目标人脸.

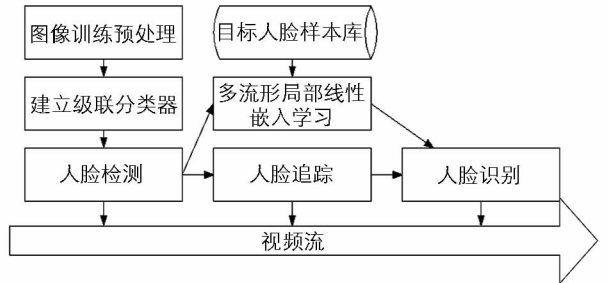


图 1 本文视频追踪算法的总体结构

2 多流形学习

流形空间与欧式空间的拓扑相似.

定义 1 流形 M 是一个度量空间, M 具有以下性质: 如果 $x \in M$, 则存在一些 x 的近邻点 U 与整数 $n \geq 0$, 使得 U 与 \mathbb{R}^n 同胚. 假设 $\epsilon > 0$ 是一个正实数, 将流形中点 x 的近邻点定义为与 x 距离小于 ϵ 的点集合 $N_\epsilon(x) = \{y \in M: d(x, y) < \epsilon\}$, 将流形 M 中每个点 x 表示为一个特征向量 $(a_1, \dots, a_n) \in \mathbb{R}^n$, 其中 $x_i \in x$ 均为实数, 将特征向量之间的距离 $d(x, y)$ 定义为马氏距离: $d(x, y) = \sum_{i=1}^n |x_i - y_i|$.

2.1 单类的单流形模型

首先考虑单类的分类问题: 使用 LLE 算法对 X 进行处理, 首先定义点 $x_i \in X$ 的 k -近邻点, 假设 $N_\epsilon(x)^c$ 表示近邻 $N_\epsilon(x)$ 的各个元素.

定义 2(k -近邻点) 假设 $\epsilon > 0$, $N_{\epsilon, k}(x)$ 表示点 $x \in X$ 的 k -近邻点, 定义为 $N_{\epsilon, k}(x) \subset N_\epsilon(x)$, $\forall x_i \in N_{\epsilon, k}(x)$, 且 $\forall x_j \in N_{\epsilon, k}(x)^c \cap N_\epsilon(x)$, $d(x, x_i) < d(x, x_j)$. 以下证明数据集中如果任意两点 x_i 与 x_j 包含相同的特征向量, 则 x_i 与 x_j 具有相同的 k -近邻点.

命题 1 假设 x_i 与 x_j 均为 \mathbb{R}^D 中的特征向量, 如果 $x_i = x_j$, 则 $N_{\epsilon, k}(x_i) \cap N_{\epsilon, k}(x_j) \neq \emptyset$.

证 对于 $x_i, x_j \in \mathbb{R}^n$, 如果 $x_i = x_j$, 则 $N_\epsilon(x_i) = N_\epsilon(x_j)$. 根据定义 2 可得 $N_{\epsilon, k}(x_i) \subset N_\epsilon(x_j)$, $N_{\epsilon, k}(x_j) \subset N_\epsilon(x_j)$, 因此 $N_{\epsilon, k}(x_i) \cap N_{\epsilon, k}(x_j) \neq \emptyset$.

2.2 多类的多流形模型

考虑多类的分类问题. 为了简化分析, 仅考虑两个类的分类问题, 假设 X 是包含 $N+P$ 个特征向量的

数据集, F 表示高维空间的数据点, 假设前 N 个点属于类 1, 后 P 个点属于类 2, 假设 $c \in \{1, 2\}$, $L \in \{N, P\}$, X 中每个特征向量 x_i^c 的维度为 D , $x_i^d = (x_{i,1}^c, x_{i,2}^c, \dots, x_{i,L}^c)^T$, $1 \leq i \leq D$.

$$F = [x_1^1 x_2^1 \dots x_N^1 x_1^2 x_2^2 \dots x_P^2] = \begin{bmatrix} x_{1,1}^1 & x_{1,2}^1 & \dots & x_{1,N}^1 & x_{1,1}^2 & x_{1,2}^2 & \dots & x_{1,P}^2 \\ x_{2,1}^1 & x_{2,2}^1 & \dots & x_{2,N}^1 & x_{2,1}^2 & x_{2,2}^2 & \dots & x_{2,P}^2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{D,1}^1 & x_{D,2}^1 & \dots & x_{D,N}^1 & x_{D,1}^2 & x_{D,2}^2 & \dots & x_{D,P}^2 \end{bmatrix}$$

传统的 LLE 算法无法同时学习多个流形, 因为 LLE 算法中决定给定点的近邻点的过程忽略了类标签. 由命题 1 知两个类的任意两点含有相同的特征向量, 这两点则具有相同的近邻点, 在流形学习中现有的算法无法保留流形的结构.

针对上述问题, 为传统的 LLE 算法设计了一个近邻点选择机制, 本方法在选择点 x_i 近邻点的过程中考虑其类标签, 将本方法称为监督 k -近邻选择. 图 2 所示, 在寻找类 1 中点 x_i 的 k -近邻点过程中, 经典方法也会考虑类 2 的点, 因此, 不属于类 1 的 x_k 为 x_i 的近邻点, 而类 1 的 x_j 与 x_i 则不会变为 x_i 的近邻点; 本文方法中 x_k 不是 x_i 的近邻点, 而 x_j 与 x_i 变为其近邻点, 所以获得一个成功的流形重构过程.

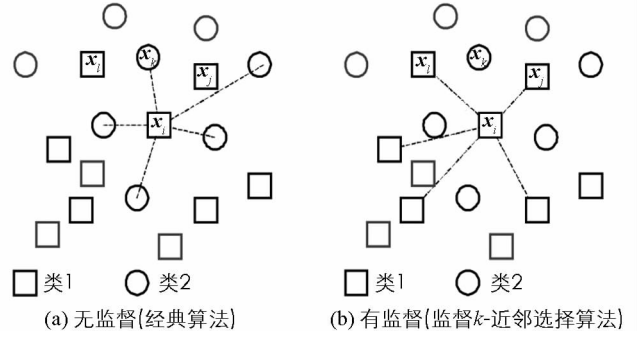


图 2 本文的近邻点选择策略

定义 3(监督 k -近邻点) 假设 C 是 X 的一个数据类, 如果点 $x_i \in C \subset X$ 的监督 k -近邻点是 $N_{\epsilon, k}(x_i)$, 则 $x_j \in C$.

定义 4(等距算子^[11]) 假设 X, Y 是线性的度量空间, $x, y \in X$, 如果 X 到 Y 的映射 f 满足条件: $d(f(x), f(y)) = d(x, y)$, 则 f 是一个等距算子, 即等距算子是一个保留了距离的映射.

选择监督 k -近邻点与每个类中数据的等距算子, 然后根据引理 1 可获得它们对应的流形.

引理 1 假设 X, Y 是线性的度量空间 $x, y \in X$, 一个等距算子 $f: X \rightarrow Y$ 将 x 的近邻映射到 Y 中 $f(x)$ 的近邻.

证 假设 $N_\epsilon(x)$ 是点 $x \in X$ 的一个近邻点集合, 假设等距算子将 x 映射到 $f(x)$, $N_\epsilon(x)$ 的每个点 y 满足 $d(x, y) < \epsilon$. 根据定义 4, 因为 $d(f(x), f(y)) = d(x, y)$, 所以每个点 $y \in N_\epsilon(x) \subset X$ 映射至 $f(y) \in N_\epsilon(f(x)) \subset Y$.

引理 2 假设 X 是线性的度量空间, C 是 X 的一个数据类, 对于任意点 $x_i \in X$,

- 1) 如果 $x_i \in C \subset X$, 则 $N_{\epsilon, k}(x_i) \subset C$;
- 2) 如果 $x_i \in C \subset X$ 等距地映射至点 $y_i \in M_1$, 则 $N_{\epsilon, k}(y_i) \subset M_1$, 其中 M_1 是 C 中所有点获得的流形.

证 根据引理 1 的定义, 将 $N_{\epsilon, k}(x_i) \subset X$ 映射到 $N_{\epsilon, k}(y_i) \subset M$, 可得 $N_{\epsilon, k}(x_i) \subset C$, 因此 $N_{\epsilon, k}(y_i) \subset M_1$. 根据引理 1 与 2 的结果可看出, 通过流形内每个点的近邻点可重构其邻域, 并保留流形的局部几何结构. 下面通过本文的近邻选择方法对降维空间的流形结构进行比较. 为了简便, 本文规定如果至少有一个点 $x_i^1 \in C_1$ 与点 $x_j^2 \in C_2$ 使得 $x_i^1 = x_j^2$, 则表示为 $C_1 \delta C_2$.

引理 3 假设 x_i^1 与 x_j^1 是有限维线性空间 X 的任意两点, C_1 与 C_2 是两个数据类, 其中 $x_i^1 \in C_1 \subset X$, $x_j^2 \in C_2 \subset X$, 如果 $x_i^1 = x_j^2$, 则 $N_{\epsilon, k}(x_i^1) \delta N_{\epsilon, k}(x_j^2)$.

证 如果 $x_i^1 = x_j^2$, 则 $N_{\epsilon, k}(x_i^1) \cap N_{\epsilon, k}(x_j^2) \neq \emptyset$, 因此 $N_{\epsilon, k}(x_i^1) \delta N_{\epsilon, k}(x_j^2)$.

理论 2 假设 $C_1, C_2 \subset X$, 即 $X = C_1 \cup C_2$, 假设 M_1 与 M_2 分别是 C_1, C_2 获得的流形, 如果 C_1 接近 C_2 , 则 M_1 接近 M_2 .

证 假设将 x_i^1 映射到点 $y_i \in M_1$, x_j^2 映射到点 $y_j \in M_2$, 则根据引理 2 可得 $N_{\epsilon, k}(y_i) \subset M_1$ 与 $N_{\epsilon, k}(y_j) \subset M_2$, 其中 C_1 映射至 M_1 , C_2 映射至 M_2 . 根据引理 3 可得 $N_{\epsilon, k}(x_i^1) \delta N_{\epsilon, k}(x_j^2)$, 根据引理 2 可将 $N_{\epsilon, k}(x_i^1)$ 映射至 $N_{\epsilon, k}(y_i)$, 将 $N_{\epsilon, k}(x_j^2)$ 映射至 $N_{\epsilon, k}(y_j)$, 所以 $N_{\epsilon, k}(y_i) \delta N_{\epsilon, k}(y_j) \Rightarrow M_1 \delta M_2$.

在分类与模式识别过程中度量流形空间的距离具有重要意义. 首先, 考虑点到流形的距离, 该距离对

于流形空间中未知数据样本的分类极为重要.

定义 5(点到流形的距离(PMD)) 假设 \mathbf{x} 是有限维向量空间 V 中的一个点, M 是一个流形, 点 \mathbf{x} 到流形 M 的距离 $d(\mathbf{x}, M)$ 定义为 $d(\mathbf{x}, M) = \min_{S_i \in M} d(\mathbf{x}, S_i) = \min_{S_i \in M} \min_{y \in S_i} \|\mathbf{x} - \mathbf{y}\|$, 式中 S_i 是 M 的局部线性空间, 因此 $M = \{S_i : i = 1, 2, \dots, m\}$, $\|\cdot\|$ 是欧氏距离.

在多流形空间中度量两个流形的距离也十分重要, 因为本算法关注于保留每个流形的几何结构, 通过测量多流形空间中两个流形的距离可比较降维空间中的流形距离, 定义 6 定义了流形-流形距离.

定义 6(流形-流形距离) 假设 M_1 与 M_2 是两个流形, 则流形-流形有向距离定义为:

$$d(M_1, M_2) = \frac{1}{T_{M_1}} \sum_{i=1}^{T_{M_1}} \min_{1 \leq j \leq T_{M_2}} \|M_1(i) - M_2(j)\|$$

式中: T_{M_1} 与 T_{M_2} 分别为 M_1 与 M_2 的长度, $M_1(i)$ 是流形 M_1 的第 i 个点. 因为 $d(M_1, M_2)$ 是有向距离, M_1 与 M_2 的无向距离计算公式为:

$$D(M_1, M_2) = d(M_1, M_2) + d(M_2, M_1)$$

2.3 基于 LLE 的多流形算法

本文基于 LLE 的多流形学习算法主要步骤如算法 1 所示.

算法 1 基于局部线性嵌入的多流形学习算法

-
1. 读取数据集 $X = \{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N\}$;
 2. FOREACH 数据分类 C_q IN X DO
 3. FOREACH 数据点 \mathbf{x}_i IN C_q DO
 4. 搜索 k_q 个近邻点 $N_{\epsilon, k_q}(\mathbf{x}_i)$;
 5. 通过最小化(1) 式计算权重 w_{ij} ;
 6. 搜索权重矩阵 \mathbf{W} , 每个元素的权重为 w_{ij} ;
 7. ENDFOR
 8. ENDFOR
 9. 搜索 $N \times N$ 稀疏矩阵 $\mathbf{F} = (\mathbf{I} - \mathbf{W})^T(\mathbf{I} - \mathbf{W})$;
 10. 设置 $d = 1, MinDist = 0$;
 11. /* 迭代地运行以下子程序直至不同分类之间的流形距离最小化 */
 12. WHILE($MinDist < \epsilon_{mnd} \& \& d < N$) DO
 13. 从 M 个特征向量中搜索 $d + 1$ 个最小特征向量构成集合 Y , 忽略最小的特征向量;
 14. /* 从 Y 中搜索流形嵌入坐标 \mathbf{Y}_q */
 15. 设置 $initialLength = 0$;
 16. FOREACH 数据类 \mathbf{X}_q IN X DO
 17. $\mathbf{Y}_q = Y((initialLength + 1), (initialLength + length(\mathbf{X}_q)))$;
 18. $initialLength = initialLength + length(\mathbf{X}_q)$;
 19. ENDFOR
 20. /* 计算流形与流形之间的距离 */
 21. FOREACH $\mathbf{Y}_q, \mathbf{Y}_r$ IN Y DO
 22. 根据定义 6 搜索 $D(\mathbf{Y}_q, \mathbf{Y}_r)$;
 23. ENDFOR
 24. /* 搜索流形之间的最小距离 */
 25. 搜索最小距离: $MinDist = \min(D(\mathbf{Y}_q, \mathbf{Y}_r))$;
 26. $d = d + 1$;
 27. ENDWHILE
-

2.4 计算复杂度分析

假设一个 D 维数据集, 大小为 n , 将 C 个数据类降维到 d 维空间, 近邻数量设为 k . 文献[12]分析了传统 LLE 算法的计算复杂度, 下文分析了本算法的计算复杂度.

本算法包含 3 个步骤:

步骤 1 每次 k -近邻搜索的平均计算成本为 $O(D \log(k) n \log(n))$, 因为本算法对各类独立地进行 k -近邻搜索, 所以平均成本为 $O(\sum_i^C D \log(k_i) n_i \log(n_i))$, 本算法在高维空间的改进明显地降低了近邻搜索的计算复杂度.

步骤 2 中权重矩阵的构建成本为 $O(\sum_i^C D n_i k_i^3)$, 而传统 LLE 的成本为 $O(D n k^3)$.

步骤 3 包含了两个子步骤: 特征值分解的成本为 $O(d n^2)$; 最小化流形-流形距离的计算成本为 $O(\sum_{i < j}^C d n_i n_j)$. 步骤 3 是一个迭代过程, 结束条件为最小化流形-流形距离达到预定的阈值, 假设算法的迭代次数为 m , 则步骤 3 的计算复杂度为 $O(m(d n^2 + \sum_{i < j}^C d n_i n_j))$.

综上所述, 本算法的总计算复杂度为:

$$O(\sum_i^C D \log(k_i) n_i \log(n_i)) + O(\sum_i^C D n_i k_i^3) + O(m d n^2 + \sum_{i < j}^C m d n_i n_j)$$

3 实验与结果分析

3.1 本文流形学习算法的分类策略

每个流形学习算法对训练数据进行流形学习, 在流形空间建立测试数据样本之后, 在低维流形空间使用 k -近邻分类方法对样本进行分类处理, 本算法的低维流形嵌入包含了多个流形. 使用点-流形的距离(定义 5)选择与给定的测试样本(\mathbf{x}_{ts})距离最近的流形, 定义为:

假设 $d(\mathbf{d}_{ts}, M_q) = \min_{M_i \in M} d(\mathbf{x}_{ts}, M_i)$, 其中 $M = [M_1, M_2, \dots, M_q, \dots, M_N]$ 是流形空间, M_q 是与类 C_q 对应的流形. $Class(\mathbf{x}_{ts}) = \begin{cases} C_q & \text{如果 } d(\mathbf{x}_{ts}, M_q) < \epsilon_{pmd} \\ \text{未决定} & \text{其他情况} \end{cases}$

3.2 视频人脸的检测与识别实验

为了测试多流形算法对人脸识别与追踪的性能, 选择两个公开的大规模视频数据集: YouTube Celebrities(YTC)^[13]与 COX^[14], YTC 广泛应用于视频人脸数据集的检测与追踪应用中, YTC 共有 47 个主题、1 910 个视频剪辑, 其中大多数剪辑经过了高度的压缩且含有噪声, 视频质量较低; COX 则是一个大规模数据集, 包含了 1 000 个主题, 每个主题包含 3 个视频(不同摄像头拍摄), 每个视频中大约有 25~175 个低分辨率的视频帧. 图 3,4 所示分别是 YTC 与 COX 数据集的部分样本.



图 3 YTC 数据集的视频帧样本



图 4 COX 数据集的视频帧样本

首先对 COX 数据集的所有人脸基于人脸边界框进行正则化处理, 获得大小接近的人脸, 采用文献[15]的方法对 YTC 数据集进行预处理: YTC 的人脸缩放为 20×20 大小, COX 的人脸缩放为 32×40 大小, 对两个数据集进行直方图均衡化处理来消除光照效应. 使用文献[15]的实验方案构建十折交叉检验, 即随机选择 10 个训练集/测试集, YTC 数据集的每折随机选择 3 个图像集作为训练集, 6 个作为测试集;

COX 数据集的每个主题有 3 个视频, 从 3 个视频测试集(每个人)中选择一个视频作为训练集, 剩下的两个视频作为测试集, 因此, 共需建立 6 个测试分组。

3.3 实验结果与分析

将本文多流形监督学习算法与无监督学习算法(MaxMD^[16], CDL^[17])、监督学习算法(SGM^[19], GMM^[20])进行比较, 全面评估本算法的检测性能。表 1 所示是视频识别算法对 YTC 数据集与 COX 数据集的实验结果, 全部结果均为十折交叉检验的平均识别率。从结果可看出, 本算法对大规模视频数据集 YTC 与 COX 的检测结果均具有明显的优势, 原因在于本算法通过多流形的学习策略有效地降低了问题的维度, 提高了目标识别的准确率, 此外, 将流形空间中流形之间的距离作为一个度量指标, 搜索数据的最优低维嵌入, 最优的低维嵌入可提高分类准确率。

表 1 视频识别算法对 YTC 数据集与 COX 数据集的识别率/%

视频检测算法	YTC	COX12	COX13	COX23	COX21	COX31	COX32
MaxMD	52.6	36.4	19.6	8.9	27.6	19.1	9.6
CDL	69.7	78.4	85.3	79.7	75.6	85.8	81.9
LMKML	70.3	66.0	71.0	56.0	74.0	68.0	60.0
SGM	52.0	26.7	14.3	12.4	26.0	19.0	10.3
GMM	61.0	30.1	24.6	13.0	28.9	31.7	18.9
本算法	73.3	95.1	96.3	94.2	92.3	95.4	94.5

表 2 所示是几种目标检测算法的计算时间统计, 可看出本算法的训练时间较低, 测试时间也较低, 具有较好的实时性。

表 2 目标检测算法的计算时间/s

视频检测算法	训练时间	测试时间
MaxMD	无训练	0.1
CDL	433.3	2.6
LMKML	245.3	0.5
SGM	11.9	0.1
GMM	42.3	1.9
本算法	27.1	0.1

4 结束语

经典的 LLE 算法难以学习多个流形, 本文设计了基于 LLE 的多流形学习方法, 通过局部非线性、多流形的方法学习并保留数据集的类结构。本算法设计了监督的近邻选择策略, 将流形空间中流形之间的距离作为一个度量指标, 搜索数据的最优低维嵌入提高分类准确率。通过寻找多流形空间中与同一数据样本距离最近的流形对已知数据进行分类, 有效地保留了每个流形的结构并保持流形之间分离性。

参考文献:

- [1] 李燕燕, 闫德勤, 刘胜蓝, 等. 一种基于局部线性嵌入的多流形学习算法 [J]. 小型微型计算机系统, 2012, 33(8): 1795-1799.
- [2] 高翠珍, 胡建龙, 李德玉. 保持局部邻域关系的增量 Hessian LLE 算法 [J]. 计算机科学, 2012, 39(4): 217-219.
- [3] 晏勇. 基于 SKLLE 和 SVM 的人脸表情识别 [J]. 西南师范大学学报(自然科学版), 2014, 01(1): 55-60.
- [4] 王鑫, 沃波海, 管秋, 等. 基于流形学习的人体动作识别 [J]. 中国图象图形学报, 2014, 19(6): 914-923.
- [5] ZHANG S Q. Enhanced Supervised Locally Linear Embedding [J]. Pattern Recognition Letters, 2009, 30(13): 1208-1218.
- [6] ZHANG W. An Improved Locally Linear Embedding Algorithm by Using Recurrent Neural Networks [J]. Journal of Computational and Theoretical Nanoscience, 2016, 13(1): 425-437.
- [7] ZHANG L, LENG Y, YANG J, et al. Supervised Locally Linear Embedding Algorithm Based on Orthogonal Matching Pursuit [J]. IET Image Processing, 2015, 9(8): 626-633.
- [8] 任剑锋, 梁雪, 李淑红. 基于非线性流形学习和支持向量机的文本分类算法 [J]. 计算机科学, 2012, 39(1): 261-263.

- [9] CHOJNACKI W, BROOKS M J. A Note on the Locally Linear Embedding Algorithm [J]. *International Journal of Pattern Recognition and Artificial Intelligence*, 2009, 23(08): 1739–1752.
- [10] 胡一帆, 胡友彬, 李 骞, 等. 基于视频监控的人脸检测跟踪识别系统研究 [J]. *计算机工程与应用*, 2016, 52(21): 1–7.
- [11] 吴文通, 李元祥, 韦邦合, 等. 局部测地距离估计的增量等距特征映射算法 [J]. *上海交通大学学报(自然科学版)*, 2013, 47(7): 1082–1086.
- [12] KAYO O. *Locally Linear Embedding Algorithm—Extensions and Applications* [M]. Oulu: Oulu University Press, 2006.
- [13] KIM M, KUMAR S, PAVLOVIC V, et al. *Face Tracking and Recognition with Visual Constraints in Real-World Videos* [C]// 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. New York: IEEE Computer Society Press, 2008.
- [14] HUANG Z, SHAN S, ZHANG H, et al. Benchmarking Still-to-Video Face Recognition via Partial and Local Linear Discriminant Analysis on COX-S2V Dataset [C]//Asian Conference on Computer Vision. New York: Springer-Verlag, 2012: 589–600.
- [15] LU J, WANG G, MOULIN P. Image Set Classification Using Holistic Multiple Order Statistics Features and Localized Multi-kernel Metric Learning [C]//IEEE International Conference on Computer Vision. New York: IEEE Computer Society Press, 2013: 329–336.
- [16] GRETTON A. A Kernel Two-Sample Test [J]. *Journal of Machine Learning Research*, 2012, 13(1): 723–773.
- [17] WANG R, GUO H, DAVIS L S, et al. Covariance Discriminative Learning: A Natural and Efficient Approach to Image Set Classification [C]//IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE Computer Society Press, 2012: 2496–2503.
- [18] MLIKI H, FENDRI E, HAMMAMI M. Face Recognition Through Different Facial Expressions [J]. *Journal of Signal Processing Systems*, 2015, 81(3): 1–14.
- [19] CUI Z, CHANG H, SHAN S, et al. Joint Sparse Representation for Video-Based Face Recognition [J]. *Neurocomputing*, 2014, 135(8): 306–312.

A Tracking Algorithm of Surveillance Video Based on Enhanced Multi-Manifold Learning

LI Jian-xin

Department of Computer Engineering, Dongguan Polytechnic, Dongguan Guangdong 523808, China

Abstract: The traditional manifold learning algorithms cannot preserve the structure of individual manifolds during multi-class-multi-manifold learning problems, and have obvious influence to the performance of multi-classes identification problems, thus a manifold learning algorithm of multi-manifold locally linear embedding has been proposed. A supervised neighborhood selection method has been designed by this multi-manifold learning algorithm for the manifold learning procedure of each class, and the distances of manifold to manifold have been set as the metric to search the optimal low dimensional space. Image training preprocess of external database has been realized during the video tracking algorithm, the cascade classifier has been constructed for face detection, and the mean particle filter combined with tracking correction strategy has been adopted for real-time tracking of face images, the results of multi-manifold learning training are used to identify the target faces from the face set of video stream. Simulation experiments are implemented based on the large scale video datasets, the results show that the proposed algorithm realizes a high detection accuracy and a high computational efficiency to different video datasets.

Key words: manifold learning; locally linear embedding; surveillance video; target detection; target tracking; dimensional reduction